

P23563

1

## Apparatus and Method for Image Recognition

### Field of the Invention

The present invention relates to an apparatus and a method  
5 for recognizing an object displayed in an input image and  
releasing data of its position and shape.

### Background of the Invention

One of conventional image recognizing apparatus is known  
10 as disclosed in the Japanese Patent of (Publication No. 9-  
21610).

Fig. 35 is a block diagram of a conventional image  
recognizing apparatus which comprises:

(a) image input unit 3511 for receiving an image of  
15 interest;

(b) model memory unit 3512 which stores local models of  
an object to be identified;

(c) matching process unit 3513 for matching each image  
segment of the input image with the local models;

20 (d) local data integrating unit 3514 for integrating and  
displaying, in probabilistic way, the position of the object  
to be identified in a parameter space together with the position  
of the image segment depending on the degree of the matching  
of each image segment of the input image with its local model;  
25 and

P23563

2

(e) object position determining unit 3515 for determining image segments with the highest probability from the parameter space to determine the position of the object to be identified in the input image.

5       The conventional image recognizing apparatus may carry out the recognizing operation with much difficulty as a number of similar local models of different models are increased.

Another conventional image recognizing apparatus is also known as disclosed in the Japanese Patent (Publication No.

10   6-215140).

Fig. 36 is a block diagram of the another conventional image recognizing apparatus which comprises:

(a) display 3601 for displaying an image;

(b) main controller 3602 for controlling operations of the  
15   entire system;

(c) internal memory 3603 for storing an operating program and the like;

(d) disk 3604 for storing a reference pattern;

(e) television camera 3605 for capturing an image of an  
20   object to be identified such as a product or a sample;

(f) image input unit 3606 for converting image data of the object captured by camera 3605 into a digital form;

(g) image rotating unit 3607 for positioning the object in a gradation image of the digital form to be faced in a given  
25   direction for each category;

P23563

3

(h) image data extracting unit 3608 for sampling the rotated image at a specific rate and extracting the gradation of each sampled image as feature data of the rotated image;

(i) dictionary generating unit 3609, having average vector calculator 3609A for calculating an average vector of the images of each category from the feature data, for determining a dictionary (a list of reference patterns) of the average vectors;

(j) identifying unit 3610 having vector distance comparator 3610A for calculating a vector of an object of an unknown category and for extracting, from the dictionary generating unit 3609, one of the average vectors which is closest to the calculated vector to identify the object on the unknown category; and

(j) parameter setting unit 3611 for optimizing the parameters for image input 3606, the image rotating unit 3607, image data extracting unit 3608, and identifying unit 3610 in each category.

The another conventional image recognizing apparatus may hardly carry out the recognizing operation in case that the images of objects which are identical in the shape but different in the gradation are grouped in one category for recognizing and classifying the objects by shape. Since similar gradation images are grouped into one category, a total number of categories increases thus requiring more time for the

P23563

4

operation.

# Summary of the Invention

A first object of the present invention is to estimate the  
5 position and the type of an object to be identified in an input  
image at high accuracy even when local models of different types  
are very similar.

An image recognizing apparatus according to the present  
invention comprises:

- 10 (a) image input means for inputting an image;
- (b) image dividing means for dividing the image received  
from the image input means into input local-segments;
- (c) similar window extracting means for extracting a  
learning-local-segment which is similar to each input  
15 local-segment received from the image dividing means;
- (d) object position estimating means for estimating a  
position of an object to be identified in the input image from  
the coordinates of the input window and the coordinates of the  
learning- window received from the similar window extracting  
20 means; and
- (e) counting means for counting a pair of the learning  
window and the input window for each position which is estimated  
from the learning window and the input window by the object  
position estimating means.

25 The operation of the image recognizing apparatus having

P23563

5

the above arrangement of the present invention includes:

(1) extracting a learning window which is similar to a input window in the input image;

(2) estimating, in the input image, a position of a model  
5 in a learning image from the coordinates of the learning window in the learning image and the coordinates of the corresponding input window in the input image; and

(3) counting a pair of the learning window and the input window for each position estimated from the learning window and  
10 the input window. Consequently, when the counted number is greater than a predetermined number, it is judged that the object of a type expressed with the learning image is present in the input image, and the position of the object can be estimated at high accuracy.

15 A second object of the present invention is to quickly recognize a shape of an object in an image and determine its position while the images of objects which are identical in the shape but different in the gradation are grouped into one category.

20 Another image recognizing apparatus according to the present invention comprises:

(a) an image database for preliminarily storing a shape identifier specifying a shape of an object to be identified and images of the object having the shape;

25 (b) model generating means for preliminarily extracting

P23563

6

feature data of the shape from the model images;

(c) a shape database for preliminarily storing the feature data of the shape with its shape identifier;

(d) an image input unit for inputting an input image to  
5 be examined;

(e) an image cutout unit for cutting out an image segment from the input image as a partial image;

(f) shape classifying means for determining whether or not the object of the shape is present in the image segment by  
10 comparing the image segment with the feature data of the shape;  
and

(g) an output unit for releasing, if there is an object having a shape which coincides with a shape in the input image, data about the shape of the object determined by the shape  
15 classifying means and about the position of the shape of the object in the input image.

The another image recognizing apparatus according to the present invention allows the feature data of the shape to be preliminarily extracted from many model images and to be  
20 compared with the input image. Accordingly, the another image recognizing apparatus can quickly examine whether or not the object is present in the input image from less amounts of data and, when so, readily provide the position and the shape of the object.

25

P23563

7

### Brief Description of the Drawings

Fig. 1 is a block diagram of an image recognizing apparatus according to Embodiment 1 of the present invention;

Fig. 2 is a block diagram of the image recognizing apparatus  
5 of Embodiment 1 implemented by a computer;

Fig. 3 is a flowchart showing a procedure in Embodiment  
1;

Fig. 4 illustrates an example of input image in Embodiment  
1;

10 Fig. 5 illustrates examples of learning image data stored  
in a learning image database of Embodiment 1;

Fig. 6 illustrates an example of a combination of an input  
window and a learning window released from similar window  
extracting means of Embodiment 1;

15 Fig. 7 illustrates an example of a resultant output of  
counting means;

Fig. 8 is a block diagram of an image recognizing apparatus  
according to Embodiment 2 of the present invention;

20 Fig. 9 is a flowchart showing a procedure in Embodiment  
2;

Fig. 10 illustrates examples of same-type images stored  
in an image database of Embodiment 2;

Fig. 11 illustrates examples of same-type window data  
stored in a same-type window database of Embodiment 2;

25 Fig. 12 illustrates an example of a combination of an input

P23563

8

window and a learning window released from similar window extracting means of Embodiment 2;

Fig. 13 illustrates an example of a resultant output of counting means of Embodiment 2;

5 Fig. 14 is a block diagram of an image recognizing apparatus according to Embodiment 3 of the present invention;

Fig. 15 is a flowchart showing a procedure in Embodiment 3;

10 Fig. 16 illustrates examples of learning image data stored in a learning image database for type X of Embodiment 3;

Fig. 17 is a block diagram of an image recognizing apparatus according to Embodiment 4 of the present invention;

Fig. 18 is a block diagram of the image recognizing apparatus of Embodiment 4 implemented by a computer;

15 Fig. 19 is a flowchart showing a procedure of the operation of model generating means of Embodiment 4;

Fig. 20 is a flowchart showing a procedure at an image input unit through an output unit in Embodiment 4;

20 Fig. 21 illustrates examples of model images and their shape identifiers stored in an image database of Embodiment 4;

Fig. 22 illustrates an example of an average model image and its shape identifier stored in a shape database of Embodiment 4;

25 Fig. 23 illustrates examples of rectangular segments cut out by an image cutout unit;



P23563

9

Fig. 24 illustrates an example of a detection result released by an output unit;

Fig. 25 is a block diagram of an image recognizing apparatus according to Embodiment 5 of the present invention;

5 Fig. 26 is a flowchart showing a procedure of the operation at model generating means of Embodiment 5;

Fig. 27 is a flowchart showing a procedure of an operation at an image input unit through an output unit in Embodiment 5;

Fig. 28 illustrates examples of model images and their shape  
10 identifiers stored in an image database of Embodiment 5;

Fig. 29 is a block diagram of an image recognizing apparatus according to Embodiment 6 of the present invention;

Fig. 30 is a flowchart showing a procedure of an operation at model generating means of Embodiment 6;

15 Fig. 31 is a flowchart showing a procedure of an operation at an image input unit through an output unit in Embodiment 6;

Fig. 32 illustrates an example of model images and its shape identifier stored in an image database of Embodiment 6;

Fig. 33 illustrates examples of rectangular segments cut  
20 out by an image cutout unit of Embodiment 6;

Fig. 34 illustrates an example of a resultant output of a counting unit of Embodiment 6;

Fig. 35 is a block diagram showing a conventional image recognizing apparatus; and

25 Fig. 36 is a block diagram showing another conventional

P23563

10

image recognizing apparatus.

### Detailed Description of the Preferred Embodiments

Exemplary embodiments of the present invention will be  
5 described in detail referring to Figs. 1 to 34.

#### (Embodiment 1)

Fig. 1 is a block diagram of an image recognizing apparatus  
according to Embodiment 1 of the present invention. Image input  
means 1 receives image data of an object to be identified. Image  
10 dividing means 2 divides the image received by image input means  
1 into input windows as local-segments. Similar window  
extracting means 3 retrieves, from a database, the data of a  
learning window which is similar to the local window from image  
dividing means 2, and releases the learning window with its  
15 corresponding local window as the learning-local-segments.  
Learning means 4 preliminarily generates an image model of the  
object to be identified. Learning image database 41 divides  
a learning image, which is a model image of an object to be  
identified, into windows having the size as the local window  
20 from image dividing means 2 and stores them as learning windows.  
Object position estimating means 5 calculates the position of  
the object in the input image from the position of the learning  
window in the learning image retrieved by similar window  
extracting means 3 and the position of the corresponding local  
25 window in the input image. Counting means 6 counts a pair of

P23563

11

the local window, which is received from object position estimating means 5, and the learning window for each position which is estimated from the window and the learning window. Object determining means 7 judges whether the object is present  
5 or not in the input image and, when so, determines the position of the object in the input image.

Fig. 2 is a block diagram showing an image recognizing apparatus implemented by a computer. The image recognizing apparatus comprises computer 201, CPU 202, memory 203, keyboard  
10 and display 204, storage medium unit 205 such as an FD, a PD, or an MO drive for holding an image recognizing program, interface (I/F) units 206, 207, and 208, CPU bus 209, camera 210 for capturing an image, image database 211 for supplying pre-stored image data, learning image database 212 for dividing  
15 the learning image, which is the model image of each object of interest, into a window-sized segment and storing them as learning windows, and output terminal 213 for delivering the type and position of the object via the I/F units.

The operation of the image recognizing apparatus having  
20 the above arrangement is now explained referring to the flowchart shown in Fig. 3. Fig. 4 illustrates an example of the input image, Fig. 5 illustrates examples of the learning image, Fig. 6 illustrates an example of the data output of similar window extracting means 3, and Fig. 7 illustrates an  
25 example of the resultant output of counting means 6.

P23563

12

In learning image database 41 (learning image database 212), the same window-size of images of an object to be identified as the input window, shown in Fig. 5, are stored as image data of a learning window together with the coordinate at the center thereof. Learning images 1 and 2 shown in Fig. 5 are provided for identifying a sedan-type vehicle in the shown direction and size.

Image input means 1, which is camera 210 or image database 211, receives an image data of interest (Step 301). Image dividing means 2 retrieves an input window data of a predetermined size from the received image through moving and locating a local window and releases the input local window data with the coordinate at the center thereof (Step 302).

Similar window extracting means 3 calculates a difference between the inputlocal window data in the input image received from image dividing means 2 and the corresponding learning window data stored in learning image database 41 (learning image database 212) (e.g. a sum of squares of a pixel data difference or an accumulation of the absolutes of a pixel data difference) and picks up one of the learning window data with the minimum difference. Picking up the most similar learning window to every input local window in the input image from learning image database 41, similar window extracting means 3 releases the coordinates at the center of the learning window and the coordinates at the center of the corresponding input local

P23563

13

window in a combination as shown in Fig. 6 (Step 303).

Object position estimating means 5, upon receiving a pair of the coordinate data of the learning window and the coordinate data of the input local window (Step 304), estimates the position of the object in the input image (more specifically, the coordinates at the upper left corner of a rectangular which circumscribes the object, i.e., at the origin of the learning image shown in Fig. 5) (Step 305). Input the coordinates ( $\alpha$ ,  $\beta$ ) of the input local window shown in Fig. 6 and the coordinates ( $\gamma$ ,  $\theta$ ) of the learning window, object position estimating means 5 releases the position of the object is expressed as ( $\alpha-\gamma$ ,  $\beta-\theta$ ).

Counting means 6, when receiving the coordinates ( $\alpha-\gamma$ ,  $\beta-\theta$ ) calculated at Step 305, increments the score for the coordinates by one (Step 306). As a procedure from Step 304 to Step 306 has been repeated for all the pairs of the input local window and the learning window (Step 307), counting means 6 releases a sum data including the coordinates at the position and the score as shown in Fig. 7.

Object image determining means 7 then judges whether or not the score for each set of the coordinates is greater than certain value T (Step 309). When so, it is judged that the object to be identified is present in the input image (Step 310). If none of the scores is greater than certain value T, it is determined that the object to be identified is not present in the input image (Step 311). The coordinates at the position

P23563

14

of the object are then passed through I/F unit 208 and released from output terminal 213 (Step 312).

(Embodiment 2)

- 5        Fig. 8 is a block diagram of an image recognizing apparatus according to Embodiment 2 of the present invention. Image input means 801 receives image data of an object to be identified. image dividing means 802 divides the image data supplied from the image input means 801 into an input window as local-segments
- 10      and releases the input window data. Similar window extracting means 803 retrieves learning window (learning-local-segment) data, which is similar to the input local window data divided by image dividing means 802, from a database and releases it together with the corresponding input local window data.
- 15      Learning means 804 preliminarily generates a model of the object to be identified. Learning image database 841 divides a learning image, which represents the model of the object to be identified, into learning windows having the same size as the input windows generated by image dividing means 802 and stores
- 20      the learning windows. Similar window integrating unit 842 makes a group of the learning windows which are stored in learning image database 841 and similar to each other and releases the image data of a representative learning window of the group together with the coordinates of each of the other
- 25      learning windows in the group. Integrating unit 842 also

P23563

15

releases the coordinates and the image data of a learning window which is dissimilar to the other learning windows. Same-type window database 843 stores the coordinates and the image data of the representative learning window of each group received  
5 from similar window integrating unit 842. Object position estimating means 805 calculates the position of the object in the input image from the position of the learning window in the learning image retrieved by similar window extracting means 803 and the position of its corresponding input local window in the  
10 input image. Counting means 806 counts a pair of the input local window and the learning window for each position which is estimated from the input window and the learning window by object position estimating means, when receiving a result of the counting operation of counting means 806, judges whether  
15 or not the object is present in the input image and, if so, determines the position of the object.

The operation of the image recognizing apparatus having the above arrangement is now explained referring to the flowchart shown in Fig. 9.

20 *Sub a1* While the input image is shown in Fig. 4 and the learning images are shown in Fig. 5, Fig. 10 illustrates similar windows stored in similar window database 841, Fig. 11 illustrates same-type window data stored in same-type window database 843, Fig. 12 illustrates a data output of similar window extracting  
25 means 803, and Fig. 13 illustrates a resultant output of

P23563

16

counting means 806.

Sub  
a2

The image of each object to be identified is divided into learning windows having the same size as the input local windows of the input image shown in Fig. 5, and each learning window data is stored together with its window number and the coordinates at the center thereof in learning image database 841. Fig. 5 shows such two learning windows as learning images 1 and 2 for identifying a sedan-type vehicle in the shown direction and size. Same-type window database 843 stores the image data of a representative learning window of each group of the similar windows, such as shown in Fig. 10, and the coordinates of each of the other learning windows in the group, which are retrieved from learning image database 841 by similar window integrating unit 842 as shown in Fig. 11.

Image input means 801 receives image data of interest (Step 901). Image dividing means 802 extracts local windows of a predetermined size from the image data as an input windows and releases their data together with the coordinates at the center thereof (Step 902).

Similar window extracting means 803 calculates a difference between the input local window received from image dividing means 802 and the representative learning window of each group stored in same-type window database 843 (e.g. a sum of the squares of a pixel data difference or an accumulation of the absolutes of a pixel data difference) and picks up a group



P23563

17

having the minimum difference from the groups. As picking up a group of the learning windows which are most similar to the corresponding input local windows, similar window extracting means 803 recognizes that all the learning windows in the group  
5 are similar (or corresponding) to the input local window. Extracting means 803 retrieves the coordinates of the representative learning local window from same-type window database 843 and releases them together with the coordinates at the center of the input window and those at the center of  
10 the learning window and the type of a vehicle attributed to the learning window as shown in Fig. 12 (Step 903).

Object position estimating means 805, upon receiving a pair of the coordinate data of the learning window and the coordinate data of the input local window (Step 904), estimates the  
15 position of the object in the input image, and more specifically, the coordinates at the upper left corner of a rectangular which circumscribes the object (i.e., the origin in the learning image shown in Fig. 5), and releases its data together with the type of a vehicle (Step 905). Upon Input the coordinates of the input  
20 local window ( $\alpha$ ,  $\beta$ ) and the coordinates of the learning window ( $\gamma$ ,  $\theta$ ) as shown in Fig. 12, object position estimating means 805 releases the position of the object ( $\alpha-\gamma$ ,  $\beta-\theta$ ).

Counting means 806, when receiving the coordinates ( $\alpha-\gamma$ ,  $\beta-\theta$ ) calculated at Step 905 together with data of the type  
25 of a vehicle, increments both the score for the coordinates and

P23563

18

the score for the type of a vehicle by one (Step 906).

It is then examined whether or not the procedure from Step 904 to Step 906 is completed for all the pairs of the input local window and the learning window (Step 907), and when so, counting  
5 means 806 delivers the coordinates of the position, the score for it, and the score for each type of a vehicle in a combination, shown in Fig. 13, to object image determining means 807.

Object image determining means 807 then determines whether or not the score for each set of the coordinates is greater than  
10 certain value T (Step 909). When so, the coordinates at each position of which score is greater than T and the type of a vehicle of which score is higher than any other scores is released (Step 910). If none of the scores is greater than  
15 certain value T, determining means 807 determines the object to be identified is not present in the input image (Step 911). The coordinates at the position and the type of the vehicle of the object are then released from the output terminal 213 through I/F unit 208 (Step 912).

### 20 (Embodiment 3)

Fig. 14 is a block diagram of an image recognizing apparatus according to Embodiment 3 of the present invention. Image input means 1401 receives image data of an object to be identified. Image dividing means 1402 divides the image data supplied from  
25 image input means 1401 into input windows as local-segments and

P23563

19

releases the input windows. Similar window extracting means  
1403 retrieves one similar learning window (learning -  
local-segment) to each input local window released by image  
dividing means 1402 from each learning database and releases  
5 it together with the corresponding input local window data.  
Learning means 1404 preliminarily generates a model of the  
object which corresponds to different categories to be  
identified. By-character learning image databases 1441,  
1442, ... divide a learning image which represents the model  
10 of the object to be identified into learning windows having the  
same size as the input windows determined by image dividing  
means 1402, and store the learning windows for each character .  
Object position estimating means 1405 calculates the position  
of the object in the input image from the position of the learning  
15 window in the learning image retrieved by similar window  
extracting means 1403 for each character and the position of  
its corresponding input local window in the input image.  
Counting means 1406 counts a pair of the input local window and  
the learning window for each position which is estimated from  
20 the input window and the learning window by object position  
estimating means 1405 for each character. Object determining  
means 1407, when receiving results of the counting operation  
of counting means 1406 for each character, judges whether or  
not the object is present in the input image, and if so,  
25 determines the position of the object.

P23563

20

The operation of the image recognizing apparatus having the above arrangement is now explained referring to the flowchart shown in Fig. 15.

An input image is shown in Fig. 4, a learning image of character 1 is shown in Fig. 5, an example of output data of similar window extracting means 1403 is shown in Fig. 6, and the learning image of character 2 is shown in Fig. 16.

In each of by-character learning image databases 1441, 1442, ... in learning means 1404, the image of the object to be identified of each character is divided into learning windows having the same size as the input windows of the input image shown in Fig. 5, and the learning windows are stored together with the window numbers and the coordinates at the center thereof. Fig. 5 shows such learning windows stored in character-1 learning image database 1441. Two learning images 1 and 2 are for identifying sedan-type vehicles in the shown direction and size. The learning images shown in Fig. 16 are stored in character 2 learning image database 1442 for identifying buses in the same location and direction as shown in Fig. 5.

Image input means 1401 receives image data of interest (Step '1501). Image dividing means 1402 extracts a input windows from the image data through moving and locating a window of predetermined size and releases the input window together with the coordinates at the center thereof (Step 1502).

P23563

21

- Similar window extracting means 1403 calculates a difference between the input local window of the input image received from image dividing means 1402 and its corresponding learning window stored in each by-character learning image database in learning means 1404 (e.g. a sum of the squares of a pixel data difference or an accumulation of the absolutes of a pixel data difference), and picks up the learning window having the minimum difference in each learning image database. Similar window extracting means 1403 picks up the most similar learning window for each input window from learning means 1404. Extracting means 1403 retrieves and releases the coordinates at the center of the learning window together with the coordinates at the center of the corresponding input window for each character as shown in Fig. 6 (Step 1503).
- Object position estimating means 1405, upon receiving a pair of the coordinate data of the learning window and the coordinate data of the input local window (Step 1504), estimates the position of the object in the input image, e.g., the coordinates at the upper left corner of the rectangular which circumscribes the object (i.e., the origin in the learning image shown in Fig. 5) (Step 1505). Upon input the coordinates ( $\alpha$ ,  $\beta$ ) of the input local window and the coordinates ( $\gamma$ ,  $\theta$ ) of the learning window as shown in Fig. 6, object position estimating means 1405 releases the position ( $\alpha$ - $\gamma$ ,  $\beta$ - $\theta$ ) of the object.
- Counting means 1406, when receiving the coordinates ( $\alpha$ - $\gamma$ ,

P23563

22

$\beta$ -0) calculated at Step 1505, increments the score for the coordinates of the window by one for each character (Step 1506).

It is then examined whether or not the procedure from Step 1504 to Step 1506 is completed for all the pairs of the input local window and the learning window for one character (Step 1507). The same procedure from Step 1504 to Step 1506 is repeated for another character. When the procedure has been completed for all the learning windows and the input local windows for all characters, counting means 1406 delivers the coordinates of the position and the score in a combination for each character shown in Fig. 7 to object image determining means 1407 (Step 1508).

Object image determining means 1407 then determines whether or not the score for each set of the coordinates is greater than certain value T (Step 1509). When object image determining means 1407 determines that the object of a particular character of which score is greater than T and higher than any other scores is present in the input image, the coordinates at the position of the object are released together with data of the character (Step 1510). If none of the scores is greater than certain value T, determining means 1407 determines that the object to be identified is not present in the input image (Step 1511). The coordinates at the position and the type of the object are released from output terminal 213 through I/F unit 208 (Step 1512).

P23563

23

## (Embodiment 4)

Fig. 17 is a block diagram of an image recognizing apparatus according to Embodiment 4 of the present invention.

5 Image database 1701 stores gradation images of objects having a common shape to be identified, and each gradation image is accompanied with a shape identifier including a shape name, a file name, and the coordinates at the upper left and the lower right corners of a rectangular which circumscribes the object

10 in an image. Model generating means 1702 retrieves all the gradation images of each shape to be identified from image database 1701 and extracts its feature. Feature level extracting unit 1721 calculates an average and a variance of each pixel in the rectangular which circumscribes the object

15 of each shape in all the gradation images received from image database 1701 and releases them together with the corresponding shape identifier. Shape database 1703 receives and stores each set of the average, the variance, and the shape identifier of each shape from feature level extracting unit 1721. Image input

20 unit 1704 inputs an image to be determined whether the object having a shape to be identified is present therein. Image cutout unit 1705 receives the shape identifier from shape database 1703, and cuts out an image segment having the same size as the shape to be identified from the input image. Shape

25 classifying means 1706 examines whether or not an object of the

P23563

24

shape to be identified is present in the image segment received from image cutout unit 1705. Segment shape classifying unit 1761 compares the image segment received from image cutout unit 1705 with a shape feature retrieved from shape database 1703 for determining that a shape in the image segment coincides with the shape feature. Output unit 1707, when receiving an output of the shape classifying means 1706 indicating that the object of the shape to be identified is present in the image segment, directs a display to display the shape and the position of the object in the input image.

Fig. 18 is a block diagram of an image recognizing apparatus implemented by a computer. The image recognizing apparatus comprises computer 1801, CPU 1802, memory 1803, keyboard and display 1804, storage medium unit 1805 such as an FD, a PD, an MO, a DVD or the like for storing an image recognizing program, interface (I/F) units 1806, 1807, and 1808, CPU bus 1809, camera 1810 for capturing an image, image database 1811 for supplying pre-stored image data, shape database 1812 storing model images of objects of various shapes together with their corresponding shape identifiers, and output terminal 1813 for delivering the shape and the position of the object identified via the I/F units.

The operation of the image recognizing apparatus having the above arrangement is now explained referring to the flowcharts shown in Figs. 19 and 20. Fig. 19 is the flowchart



P23563

25

showing an operation of the model generating means. Fig. 20 is the flowchart showing a procedure from inputting an image data to be examined to outputting a result of recognition. Fig. 21 illustrates examples of the model image stored in image database 1701. Fig. 22 illustrates an example of the average data of one shape with its shape identifier stored in shape database 1703. Fig. 23 illustrates an input image received from image input unit 1704 and includes rectangular image segments cut out by image cutout unit 1705. Fig. 24 illustrates a resultant output of shape classifying means 1706 displayed on the display with output unit 1707.

Prior to recognition, data about the shapes to be identified are prepared. Image database 1701 stores gradation images of various objects such as shown in Fig. 21 in the form of files as a model image. Each image is accompanied with a shape identifier including the shape image, an image file name, and the coordinates at the upper left and the lower right corners of a rectangular which circumscribes the object as an object area. The model images shown in Fig. 21 illustrate different sedan-type vehicles captured from the common angle and distance by a camera.

When a "sedan A"-type vehicle such as shown in Fig. 21 is an object of a shape to be identified, model generating means 1702 retrieves all model images accompanied with a shape identifier including the shape name of "sedan A" from image

P23563

26

database 1701 together with the shape identifier. Then,  
feature level extracting unit 1721 calculates an average image  
of rectangular sized images determined as objective areas (Step  
1901). As the model images carry objects of the same shape,  
5 their cutout image segments as the objective areas are equal  
in the size and the average image is also identical in the size.

The average image of "sedan A" shown in Fig. 21 consists  
of 148 pixels in horizontal by 88 pixel in vertical. Then,  
feature level extracting unit 1721 calculates a variance from  
10 the pixel in the rectangular as the objective area and the  
corresponding pixel of the average image for each model image  
(Step 1902).

Finally, feature level extracting unit 1721 releases the  
average image of "sedan A", the variance for each pixel, and  
15 the corresponding shape identifier, then, shape database 1703  
stores them (Step 1903). In case that a plurality of objects  
of shapes to be identified are provided, the procedure from Step  
1901 to Step 1903 is repeated for examining the respective  
shapes.

20 For recognition of the "sedan A"-type vehicle, image input  
unit 1704 (camera 210 or image database 211) supplies an input  
image (Step 2001). Image cutout unit 1705 cuts out, from the  
input image, each image segment which is equal in the size to  
the average image of "sedan A" stored in shape database 1703  
25 through moving the rectangular window having the same size as

P23563

27

the average image as shown in Fig. 23 (Step 2002).

Shape classifying means 1706 receives one image segment from image cutout unit 1705 and the average image of "sedan A" and the variance from shape database 1703. Segment shape

5 classifying unit 1761 calculates the square of a difference between each pixel of the image segment and the corresponding pixel of the average image, divides the square by the variance, and calculates a sum of the quotients to determine the distance between the image segment and the average image (Step 2003).

10 In case that the objects of shapes to be identified are two or more, segment shape classifying unit 1761 repeats the operation of Step 2003 for each shape (Step 2004). When the least calculated distances is less than a certain value (Step 2005), it is judged that the image segment contains an object of the

15 shape of the average image which is pertinent to the least distance (Step 2006).

Segment shape classifying unit 1761 judges that no object is present in the image segment when the least distance is not less than the certain value (Step 2007). The above operation

20 is repeated by segment shape classifying unit 1761 for each segment image separated from the input image (Step 2008). When it is judged that the image segment contains the object, output unit 1707 places the shape of the object over the segment image in the input image as shown in Fig. 24 (Step 2009). A resultant

25 image is then released via I/F unit 1808 from output terminal

P23563

28

1813.

## (Embodiment 5)

Fig. 25 is a block diagram of an image recognizing apparatus  
5 according to Embodiment 5 of the present invention.

Image database 2501 stores gradation images of various  
objects of shapes to be identified. Database 2501 also stores  
a shape identifier specifying the name of each shape, the image  
file name, and the coordinates at the upper left and the lower  
10 right corners of a rectangular of a predetermined size which  
circumscribes the object of the shape to be identified. Model  
generating means 2502 retrieves all the gradation images of each  
shape of the object to be identified from image database 2501  
and extracts the feature of the images. Feature space  
15 generating unit 2520 generates a feature space from the model  
images received from image database 2501 and transfers its base  
vector to shape database 2503 where each model image is  
projected to the feature space as a model image vector. Feature  
level extracting unit 2521 calculates an average and a variance  
20 of all the model image vectors of the shapes received from  
feature space generating unit 2502 for each shape and releases  
them together with the relevant shape identifier. Shape  
database 2503 receives and stores the base vector in the feature  
space from feature space generating unit 2520 and the average  
25 and variance of the model image vectors of each shape with the

P23563

29

shape identifier from feature level extracting unit 2521. Image input unit 2504 supplies an image to be determined whether the object of a shape to be identified is present therein. Image cutout unit 2505 is responsive to the shape identifier from

5 shape database 2503 for cutting out a segment image, which is equal in the size to the shape to be identified, from the input image. Shape classifying means 2506 examines whether or not an object of the shape to be identified is present in the image segment received from image cutout unit 2505. Feature space

10 projecting unit 2560 projects, to the feature space, the image segment received from image cutout unit 2505 as an image segment vector based on the base vector received from shape database 2503. Segment shape classifying unit 2561 calculates a distance between the segment image vector received from feature

15 space projecting unit 2560 and the average of model shape vectors retrieved from shape database 2503, and classifying unit 2561 determines whether or not the image segment coincides the shape to be identified. Output unit 2507, when receiving an output of the shape classifying means 2506 indicating that

20 the object of the shape to be identified is present in the image segment, display the shape and the position of the object in the image segment on a display.

The operation of the image recognizing apparatus having the above arrangement is now explained referring to the

25 flowcharts shown in Figs. 26 and 27. Fig. 26 is the flowchart

P23563

30

showing an operation of the model generating means. Fig. 27 is the flowchart showing a procedure from inputting an image to be examined through outputting a result of recognition. Fig. 28 illustrates examples of the model image stored in image database 2501. Fig. 23 illustrates an example of a rectangular segment image which is cut out by image cutout unit 2505 from an input image received from image input unit 2504. Fig. 24 illustrates a resultant output of shape classifying means 2506 displayed on the display.

10 Prior to recognition, databases about the shapes to be identified are prepared. Image database 2501 stores gradation images of various objects in the form of files such as model images as shown in Fig. 28. Each image is accompanied with the shape identifier specifying a shape of an object, an image file name, and the coordinates at the upper left and the lower right corners of a rectangular which circumscribes the object as an image area. The model images shown in Fig. 28 show a sedan-type vehicle and a bus captured from the common angle and distance by a camera.

20 When the "sedan A"-type vehicle and the bus shown in Fig. 28 are the objects of the shapes to be identified, model generating means 2502 retrieves all the model images of vehicles accompanied with the shape identifier specifying the shape name of "sedan A" and all the model images accompanied with the shape identifier specifying the shape name of "bus rear portion" from

P23563

31

image database 2501 and transfers them to feature space generating unit 2520 together with the shape identifiers.

Feature space generating unit 2520 calculates an eigenvalue and an eigenvector from the pixel in the rectangular area in the model image (Step 2601). The rectangular areas in each model image are equal in the size, each consisting of 148 pixels in horizontal by 88 pixels in vertical as shown in Fig. 28. For the bus, the rectangular shape shown in Fig. 28 is an objective area at the same position for all the model images of buses. A vector formed as a row of all the pixels in each model image is generated, and an average vector of the vectors is calculated and subtracted from each vector for determining the eigenvalue and an eigenspace.

Feature space generating unit 2520 stores the eigenvectors corresponding to the N greatest eigenvalues as base vector in shape database 2503 (Step 2602). Using the N eigenvalues, generating unit 2520 projects each model image as a model image vector in the feature space (Step 2603).

Feature level extracting unit 2521 receives the model image vector with its shape identifier from feature space generating unit 2520 and calculates an average and a covariance of the model image vectors having the same shape identifier (Step 2604). Feature level extracting unit 2521 releases the average of the model images and the average and the covariance of model image vectors of each shape together with their corresponding shape

P23563

32

identifiers , and shape database 2503 stores them (Step 2605).

For the recognition, an image to be identified is supplied from image input unit 2504 (Step 2701). Image cutout unit 2505 determines the size of a mode image from the objective area  
5 specified by the shape identifier stored in shape database 2503. Then, image cutout unit 2505 cuts out image segments having the same size through moving a window from the input image as shown in Fig. 23 (Step 2702).

Shape classifying means 2506 receives one image segment  
10 from image cutout unit 2505 and the base vector from shape database 2503. Feature space projecting unit 1760 projects the image segment as a image segment vector in the eigenspace (Step 2703). Segment shape classifying unit 2561 receives the image segment vector from feature space projecting unit 2560 and the  
15 average vectors and covariances of the "sedan A"-type vehicle and the bus from shape database 2503 respectively, and calculates a Mahalanobis distance between the image segment vector and the average vector (Step 2704).

When the least Mahalanobis distances is less than a certain  
20 value (Step 2705), it is judged that the image segment contains an object of the shape pertinent to the average vector of the least distance (Step 2706). If the least distance is not less than the certain value, it is judged that the image segment contains no object to be identified (Step 2707). Feature space  
25 projecting unit 2560 and segment shape classifying unit 2561



P23563

33

repeats the process from Step 2703 to Step 2707 for each of the image segments which are cut out from the input image (Step 2708). When it is judged that the image segment contains the object, output unit 2507 places the shape of the object over the image segment in the input image as shown in Fig. 24 (Step 2709). A resultant image is then released via I/F unit 1808 from output terminal 1813.

(Embodiment 6)

10 Fig. 29 is a block diagram of an image recognizing apparatus according to Embodiment 6 of the present invention.

Image database 2901 divides each of gradation images of various objects of the shape to be identified into rectangular shape segments having a predetermined size and stores each of the shape segments with the shape identifier specifying a shape name, a file name, and coordinates at the upper left and the lower right corners of the shape segment. Model generating means 2902 retrieves all the gradation images of the object of the shape to be identified from image database 2901 and extracts its feature. Feature space generating unit 2920 generates a feature space from the pixel values of all the shape segments in each model image received from image database 2901 and transfers its base vector to shape database 2903 where each shape segment is projected as a model image local vector to the feature space. Feature level extracting unit 2921 calculates

P23563

34

an average and variance of all the model image local vectors received from feature space generating unit 2902 for each shape segment and releases them together with the relevant shape identifier. Shape database 2903 receives the base vector of

5 the feature space from feature space generating unit 2920 and the average and variance of the model image local vectors together with the shape identifier for each shape segment from feature level extracting unit 2921, and stores them. Image input unit 2904 supplies an input image to be determined whether

10 an object of a shape to be identified is present therein. Image cutout unit 2905 is responsive to a shape identifier from shape database 2903 and cuts out an image segment having the same size as the shape segment from the input image. Shape classifying means 2906 examines whether or not an object of a shape to be

15 identified is present in the image segment received from image cutout unit 2905. Feature space projecting unit 2960 projects the image segment received from image cutout unit 2905 to the feature space as an image segment vector based on the base vector received from shape database 2903. Segment shape classifying

20 unit 2961 calculates a distance between the image segment vector received from feature space projecting unit 2960 and each average of model image local vectors retrieved from shape database 2903 and determines whether or not the image segment vector matches the shape segment of the shape of the object to

25 be identified. As shape segment classifying means 2961 detects

P23563

35

the shape segment of the shape of the object to be identified, overall shape area estimating unit 2962 estimates the area in which the overall shape of the object exists in the input image from the position of the shape segment in relation to the overall

5 shape. Counting unit 2963 counts the position of the overall shape of the object received from the overall shape area estimating unit 2962 for each image segment containing the shape segment of the shape of the object. Upon judging that the object is located at the position which is determined a number of times

10 greater than a certain number by counting unit 2963, output unit 2907 displays the shape and the position of the object on a display.

The operation of the image recognizing apparatus having the above arrangement is now explained referring to the

15 flowcharts shown in Figs. 30 and 31. Fig. 30 is the flowchart showing an operation of the model generating means. Fig. 31 is the flowchart showing a procedure from inputting an image to be examined through outputting a result of recognition. Fig. 32 illustrates examples of the model image stored in image

20 database 2901. Fig. 33 illustrates an input image which is received from image input unit 2904 and includes rectangular image segments cut out by image cutout unit 2905. Fig. 34 illustrates a resultant output of counting unit 2963. The resultant output of shape classifying means 2906 to be displayed

25 on the display is shown in Fig. 24.

P23563

36

Prior to the recognition, data about the shapes to be identified are prepared. Image database 2901 stores gradation images of the object in the form of files such as shown in Fig. 32. Each gradation image is divided into local-segments having a predetermined size of a rectangular and each local-segment is accompanied with the shape identifier specifying a local-segment name, a file name, and the coordinates at the upper left and the lower right corners of the local-segment. The local-segment name comprises the name of an overall shape of "sedan A" of the object, and a number identifying the local-segment in the overall shape. The number denotes the same position regardless of the overall shape of the object to be identified. The local-segments may overlap each other. Fig. 32 shows an example of a model image for identifying a sedan-type vehicle captured by a camera from the shown angle and distance. Actually, local-segments of plural images of similar looking sedan-type vehicles are stored together with shape identifiers of the local-segments.

When the "sedan A"-type vehicle such as shown in Fig. 32 is an object to be identified, model generating means 2902 retrieves all the model images of vehicles accompanied with the shape identifier of "sedan A" from image database 2901 and transfers them with the shape identifiers to feature space generating unit 2920. Generating unit 2920 calculates an eigenvalue and an eigenvector from the pixel value in each

P23563

37

rectangular local-segment in the model image accompanied with the shape identifier as a local model image (Step 3001). The local model images in each model image are equal in the size, each consisting of 29 pixels in horizontal by 22 pixels in vertical as shown in Fig. 32. To calculate the eigenvector, a vector which is a row of the pixels in each local model image is generated, and an average of the vectors is calculated and subtracted from each vector for determining the eigenvalue and an eigenspace.

10 Feature space generating unit 2920 generates the eigenvector corresponding to N greatest eigenvalues as the base vector, and shape database 2903 stores it (Step 3002). Using the N eigenvalues, feature space generating unit 2920 projects each local model image in the feature space to generate a local  
15 model image vector (Step 3003). Feature level extracting unit 2921 receives the local model image vector with its shape identifier from feature space generating unit 2920 and calculates an average and covariance of the local model image vectors accompanied with the same shape identifier (Step 3004).

20 Feature level extracting unit 2921 releases the average of all the local model images and the average and covariance of local model vectors for each shape, and shape database 2903 stores them together with their corresponding shape identifiers (Step 3005).

25 For recognition, an image to be determined whether an object

P23563

38

to be identified is present therein is supplied from image input unit 2904 (Step 3101). Image cutout unit 2905 calculates the size of a local-segment from the objective area determined by the shape identifier stored in shape database 2903. Then, image cutout unit 2905 cuts out each segment having the same size from the input image through moving a window as shown in Fig. 33 (Step 2702).

Shape classifying means 2906 receives one image segment from image cutout unit 2905 and the base vector from shape database 2903. Feature space projecting unit 2960 projects the image segment in the feature space as a partial image vector (Step 3103). Segment shape classifying unit 2961 receives the image segment vector from feature space projecting unit 2960 and the average vectors and covariances of local-segments of "sedan A" from shape database 2903 to calculate a Mahalanobis distance between the image segment vector and each average vectors (Step 3104). Then, classifying unit 2961 releases the shape identifier belonging to the local-segment having the average vector pertinent to the least distance.

Overall shape estimating unit 2962 calculates a difference between the coordinates at the upper left corner of the objective area defined by the shape identifier and the coordinates at the upper left corner of the image segment in the input image, and counting unit 2961 increments the score for the coordinates by one (Step 3105). The coordinates for

P23563

39

which score is incremented represent the position of the object in the input image.

Shape classifying means 2906 including feature space projecting unit 2960 through counting unit 2963 repeats the process from Step 3103 to Step 3105 for all the image segments which are cut out from the input image (Step 3106). The result in counting unit 2963 is shown in Fig. 34. In Fig. 34, a series of the coordinates are listed with the score for them in the order from the highest score. When the score for the coordinates is greater than a certain number (Step 3107), it is judged that the object is present at the coordinates in the input image (Step 3108). Then, output unit 2907 places the shape of the object over the image segment in the input image as shown in Fig. 24 (Step 3110). If none of the scores is greater than the certain number at Step 3107, it is then judged that the input image carries no object to be identified (Step 3109), and the input image is directly released (Step 3110). A resultant image is then released via I/F unit 1808 from output terminal 1813.

As set forth above, the object recognizing apparatuses of the present invention can readily detect a feature of a shape of objects from a less amount of model data even if the surface color of the object is different. Also, even if an object to be identified is partially visible in an input image, the apparatuses of the present invention can detect its shape and

**P23563**

40

position can be detected at higher accuracy.